10-19-00

A

jc956 U.S. PTO
10/17/00

UTILITY PATENT APPLICATION TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

jc893 U.S. PTO
09/691419
10/17/00

Attorney Docket No.: 4103-61641

Inventors: Gordon MacKay of 1700 Halford Avenue, Apt. 202, Santa Clara, California 95051
James P. Rivers of 14760 Live Oak Lane, Saratoga, California 95070
Rita Ousterhout of 726 Ashby Drive, Palo Alto, California 94301
Sean X. Wang of 40784 Laguna Place, Fremont, California 94539
John K. Chen of 260 North Mathilda, #M8, Sunnyvale, CA 94086

Express Mail Label No.: **EL617197949US**

Title: METHOD AND APPARATUS TO DETECT AND BREAK LOOP CONFIGURATION

Assistant Commissioner for Patents
Box Patent Application
Washington, DC 20231

Enclosed for filing with the above-identified utility patent application, please find the following:

1.	[X]	Specification (Total Pages of Text, including Abstract and Claims: 25)
2.	[X]	Drawing(s) (35 USC 113) (Total Sheets: 7) [] FORMAL [X] INFORMAL
3.	[]	Oath or Declaration (Total Pages:) [] Signed [] Unsigned
4.	[]	Microfiche Computer Program (Appendix)
5.	[]	Nucleotide and/or Amino Acid Sequence Submission (if applicable, all necessary)
a.	[]	Computer Readable Copy
b.	[]	Paper Copy (identical to computer copy)
c.	[]	Attorney for applicants hereby asserts pursuant to 37 CFR § 1.821(f) that the content of the paper of computer readable copies of SEQ ID No:1 through SEQ ID No: submitted herewith are identical
6.	[]	Assignment Papers *(cover sheet & document(s))*
7.	[]	37 CFR 3.73(b) Statement *(when there is an assignee)*
8.	[]	Power of Attorney
9.	[]	English Translation Document *(if applicable)*
10.	[]	Information Disclosure Statement (IDS/PTO-1449)
11.	[]	Copies of IDS Citations (Number of References:)
12.	[]	Preliminary Amendment
13.	[X]	Return Postcard (MPEP 503) *(should be specifically itemized)*
14.	[]	Small Entity Statement(s)
15.	[]	Certified copy of Priority Document(s)
16.	[X]	**NO FEE IS ENCLOSED AT THIS TIME**
17.	[]	Other:

"EXPRESS MAIL" MAILING LABEL NUMBER. EL617197949US
DATE OF DEPOSIT October 17, 2000

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1 10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D.C. 20231

TYPED OR PRINTED NAME Maisie C. Livengood

SIGNATURE: Maisie C Livengood

FEE CALCULATION:

	(COL. 1) NO. FILED			(COL. 2*) NO. EXTRA	SMALL ENTITY RATE	SMALL ENTITY FEE		LARGE ENTITY RATE	LARGE ENTITY FEE
BASIC FEE:						$355.00	OR		$710.00
TOTAL CLAIMS:	21	-	20	1	X $9 =		OR	X $18 =	$18.00
INDEP. CLAIMS:	5	-	3	2	X $40 =		OR	X $80 =	$160.00
____ MULTIPLE DEPENDENT CLAIMS					+ $135 =		OR	+$270 =	$0.00
*IF THE DIFFERENCE IN COL. 2 IS LESS THAN ZERO, ENTER "0" IN COL. 2.					TOTAL:				$888.00

OTHER INFORMATION:

1. [] The Commissioner is hereby authorized to debit any underpayments or credit any overpayment to Deposit Account No. 19-1970.
2. [] The Commissioner is hereby authorized to charge all required fees for extensions of time under §1.17 to Deposit Account No. 19-1970.
3. [] Foreign Priority benefits are claimed under 35 USC §119 of Patent Application Serial No. filed
4. Correspondence Address:

Richard L. Hughes
SHERIDAN ROSS P.C.
1560 Broadway, Suite 1200
Denver, Colorado 80202-5141
Telephone: (303) 863-9700
Facsimile: (303) 863-0223

Respectfully submitted,

SHERIDAN ROSS P.C.

Richard L. Hughes
Registration No. 31,264

Date: oct. 16, 2000

Attorney Docket: 4103-61641

METHOD AND APPARATUS TO DETECT AND BREAK LOOP CONFIGURATION

Inventors: Gordon MacKay
James P. Rivers
Rita Ousterhout
Sean X. Wang
John K. Chen

Assignee: Cisco Technology Inc.

CISCO-6030

Sheridan Ross P.C.
Suite 1200
1560 Broadway
Denver, CO 80202-5141

METHOD AND APPARATUS TO DETECT AND BREAK LOOP CONFIGURATION

Cross reference is made to U.S. Patent Application Serial No. 09/321,066, filed May 27, 1999, entitled Distributed Network Repeater System (Attorney File No. 4103-49601); Serial No. Serial No.: 09/330,434 filed June 11, 1999, entitled Closely-Positioned Multiple GBIC Connector (Attorney File No. 4103-29087) and Serial No.: 09/330,733, filed June 11, 1999, entitled Distributed Network Repeater Module and Method (Attorney File No. 4103-50321) all incorporated herein by reference.

The present invention relates to a method and apparatus which can automatically detect the presence of a loop configuration and can automatically break the loop, in a portion of a network or other electronic system, and in particular to a system which can reliably select a component as a loop breaking master.

BACKGROUND INFORMATION

A number of electronic and/or optical systems can be configured to provide a plurality of nodes with the nodes communicating among one another (and/or with other parts of the system) over a plurality of communication links. Although some or all features of the present invention can be applied to substantially any electrical, optical or electro-optical system having a plurality of nodes communicating over links, one useful illustrative example involves a plurality of nodes, each of which is a repeater for use in the context of a network transceiver, such as an Ethernet transceiver or switch. Although some or all features of the present invention can be used with any of a plurality of communication links (such as optical fiber links, infrared (IR) radio or other wireless links and the like), in one illustrative example, the links can include cables connecting ports of the repeaters to one another.

In any group of connected nodes, two classes of connection topologies are possible: open and closed topologies. In open topologies (a message sent from a first node cannot return to the first node without passing through at least one of the communication links more than once. Fig. 1A illustrates one example of an open-class topology, in this case a linear arrangement. Fig. 1B

illustrates a closed-class or "loop" topology, in which it is possible for a message sent from a first node to return to the same node without traversing any communication link more than once. For example, in the open system of Fig. 1A, if a message is sent from the top node to the bottom node (following the port path: A_1, B_1, A_2, B_2, A_N, it is impossible for the message to be returned to the top node without passing a second time through at least one of the communication links 114a, 114b, 114n. In contrast, in Fig. 1B, the communication links 114a, 114b, 114n, 114p form a closed or loop configuration. A message sent from the top node to the bottom node via communication links 114a, 114b, 114n can return to the top node via communication link 114p, i.e. without traversing, a second time, any of the communication links.

Although there may be many multi-node electrical or optical systems in which either (or both) of a closed-class topology or an open-class topology may be used, there are also some systems in which it is desired to avoid or eliminate loop configurations. One example is when the nodes are repeaters of an Ethernet transceiver. In this example, the presence of a loop configuration among nodes can result in collisions of packets or other communications (i.e. the presence of two or more packets on the same link or node during substantially the same time period). Accordingly, it would be useful to provide a system which can detect the presence of a loop configuration. It would further be useful to provide a system which can, preferably substantially automatically (i.e. without the need for human control, or manipulation), break the loop or otherwise reconfigure the system to eliminate the loop configuration (preferably converting it to an open-topology configuration).

Many electrical or optical systems operate according to one or more communication protocols, e.g. defining items such as the size and fields of communication packets (if any), the steps to be taken in response to certain types or contents of packets and the like. Because it can be disruptive and expensive to redesign and implement a new communication protocol, especially for systems that already have a relatively large installed base of apparatus, it would be useful to provide a system, a method and apparatus for detecting and/or breaking loop configurations which is substantially compatible with at least some existing communication protocols in the sense of avoiding substantially interfering with communication protocols used by a current installed base of apparatus. Preferably, a system, method and apparatus to detect and/or

break configurations can operate quickly (preferably requiring less than about 15 seconds, more preferably less than about 10 seconds and even more preferably less than about 5 seconds) to perform detection and/or loop breaking operations e.g. in a system of 8 nodes or less. Preferably, such a system is reliable, such as being substantially immune to at least certain types of communication errors or losses and/or without producing undefined states, and is preferably relatively easy and/or inexpensive to implement, such as requiring only (or, in some embodiments, mostly) software changes in order to implement an existing apparatus.

In some electrical or optical multi-node systems, the effective loss of a node and/or a communication link (e.g. from hardware or software failure, environmental challenge, operator error or the like) can disrupt the system such as by isolating one or more nodes in a group from communicating with other nodes in the group. For example, in the configuration depicted in Fig. 1A, if the last communication link 114n becomes inoperative, the last node 112c cannot communicate with the remaining nodes 112a 112b. Accordingly, it would be useful to provide an electrical or optical system in which the effective loss of a communication link or node, or other isolation of a node (or group of nodes) can be detected. It would further be advantageous to provide a system in which, in response to at least some types of isolation of one or more nodes, a communication link or path to the isolated node or nodes can be reestablished, preferably substantially automatically.

SUMMARY OF THE INVENTION

The present invention includes a recognition of the existence, source and/or nature of certain problems, including as described herein. In one aspect, the presence of a loop is detected by a procedure which involves a node sending a communication to one or both of its neighbors, each neighbor, in turn, passing the communication on to the next neighbor. The communication includes a value or characteristic with the property that only one of the nodes has (or is associated with) the particular value or characteristic. Each node compares its own value or characteristic with that contained in the received message and substitutes its own value or characteristic into the message (before transmitting it to its neighbor) only if its own value or characteristic is closer to the particular value or characteristic than the value or characteristic which was contained in the

received communication. As one illustrative example, the value or characteristic can be the node address and the particular value or characteristic can be the lowest (or highest) node address in the system. In this example, a node begins the process by transmitting a loop-detecting message to, e.g., one of its neighbors with the initiating node placing its own address in the message. The next node, upon receiving the message compares the received address to its own address and substitutes its own address if its own address happens to be smaller than the received address (i.e. happens to be closer to the lowest node address in the system). After such substitution (if any) the message is then passed on to the next node (if there is a next node) which performs a similar process, substituting its own address only if it is lower than the received address. Each node also, in doing the comparison, can detect if the received address is equal to its own address. In the described example, a node can only receive a message having an address equal to its own address if (a) the node has the lowest address of any node in the actively communicating system and (b) the node had previously sent out its own address in a loop-detecting communication (since this is the only way for this address to be placed into a circulating loop-detecting message). However, at this point, it is known that a node has sent out a message and the message has passed through the communication system, without passing through any link or node more than once, and yet has returned to the originating node. In other words, at this point it is known that the nodes and communication links form a closed or loop configuration.

Although the system can be initiated in any of a number fashions, including on a periodic basis, in response to a request from an external controller or other source, preferably a loop detection process is initiated in response to detecting that a new node and/or communication link has been added or deleted from the system.

In one aspect, the system provides the ability to, preferably substantially automatically, reconfigure a closed system to place it into an open-type topology. In one embodiment, the invention is implemented in a system in which a loop configuration is substantially circular (each node has exactly two communication links and each communication link is coupled to exactly two nodes). In at least this type of system, the loop can be broken by effectively disabling any single link, and the system will still be operable as an open system, with each node still being able to, at least indirectly, communicate with each other node. In order for a successful

reconfiguration to occur (effectively disabling a single link, but no more than one link) the present invention preferably provides a system, method and apparatus which provides for a single one of the nodes to act as a loop breaking master. Although it would be possible to pre-designate one of the nodes as the master (such as providing it with a unique hardware or other signature or capability) such an approach may not be as desirable as other approaches. Predesignating a node (or communication link) as a master (or otherwise unique) places a burden on the user to assure that every system is installed in such a manner as to have exactly one (and no more than one) unique node (or communication link). This burden may be infeasible, especially in systems in which nodes or communication links are substantially modular (can be readily added or removed by users). Additionally, such a system would typically fail to detect and/or break loops if the unique node or communication link became disabled or was removed.

According to one embodiment of the invention, the same system which is used for detecting a loop is also used for designating the loop-breaking master node. In one embodiment, whenever a node detects a loop (by receiving a loop detection message which contains the node's own address) that node designates itself as a loop-breaking master node. Of course, it is possible to provide numerous variations such as always designating the next-higher-address neighbor as the master, and the like. It is also possible for the master node to delegate some or all operations involved in breaking the loop to a different node.

For example, if the first node 112a in the illustration of Fig. 1B is the master loop breaker, the first node 112a can break the loop by deactivating a communication link, such as the link connected to its second or "B" port, 114a, as illustrated in Fig. 1C (with a deactivated communication link being illustrated in phantom). Preferably, deactivation of a link is provided in a reversible manner, i.e., so that the deactivated link can later be re-activated, including as described below. For example, link 114a can be deactivated by operating a switch (preferably an electronic switch) which prevents signals passing through the "B" port. It would also be possible to effectively deactivate communication link 114a using a substantially software procedure, such as causing the first node 112a to suspend the sending of any messages out through the "B" port and ignoring any messages received at the B port.

In one aspect, embodiments of the present invention can provide for reactivating a currently-deactivated communication link, e.g., in response to detecting a condition in which one or more nodes are isolated (cannot communicate with at least some other nodes). As one illustrative example, if the system as illustrated in Fig. 1C suffered a change which caused communication link 114b to become substantially inoperable, as depicted in Fig. 1D (e.g. from a hardware failure, software failure or operator action), then the second node 112b is isolated. The second node 112b cannot communicate via the A port because communication link 114a was previously (intentionally) deactivated, and cannot communicate via the "B port because communication link 114b has become substantially inoperable, as described in this example. In at least some embodiments of the present invention, in response to detecting this condition (e.g. via a "link lost" condition detected at the third node 112c and communicated to the first node 112a via communication link 114p) one of the nodes, preferably the same node which originally (intentionally) deactivated link 114a, will reactive link 114a, providing the configuration depicted in Fig. 1E. As can be seen in Fig. 1E, the system no longer contains isolated nodes (node 112b can communicate with node 112a via link 114a) and the system does not contain a loop (since communication link 114b is substantially inoperable).

As illustrated in Figs. 1C-1E, it can be advantageous to provide a system which includes one or more controllable (activatable and/or deactivatable) redundant communication links (even though this may require judicious control or configuration to avoid loops) since controllable redundant communication links can be activated or otherwise used to maintain desired communication capabilities even after there has been a failure (such as a loss of a communication link). Accordingly, in one embodiment of the invention, systems are intentionally configured (and/or users are instructed to configure their systems) such that one or more redundant links are provided, with the system (preferably substantially automatically) reconfiguring the system to deactivate links, as needed to avoid loops, while preferably retaining the capability to (preferably substantially automatically) reactivate links in order to maintain desired communication (e.g. in response to a change which isolates one or more nodes).

In one aspect, in a system having a plurality of nodes (such as Ethernet repeaters), coupled by communication links (such as cables), systems and protocols are provided for

detecting and/or breaking loops. In one aspect, in response to an added link, a repeater sends a "detect loop" message, containing its own address, to at least one neighbor. Each repeater which receives the "detect loop" message, in turn, sends it to its own neighbor, with the lesser of the received address and its own address. A repeater which receives a "detect loop" message containing its own address declares itself a master loop-breaker and can isolate one of its ports to break the loop. In one aspect, a previously intentionally-isolated port can be re-activated, e.g., in response to the loss of a communication link which could potentially isolate one or more nodes.

BRIEF DESCRIPTION OF THE DRAWINGS

Figs. 1A through 1E are block diagrams of a system of nodes and communication links in various states, including states that can be achieved according to embodiments of the present invention;

Fig. 2 is a block diagram of an Ethernet switch and coupled repeaters of a type which can be used in connection with implementing embodiments of the present invention;

Fig. 3 is a flow chart of a process for initiating a loop detect procedure;

Fig. 4 is a flow chart of a process providing loop detect and breaking according to an embodiment of the present invention;

Figs. 5A through 5E are block diagrams of a plurality of repeater modules and selected communications therebetween at various stages during a loop detect process according to an embodiment of the present invention;

Figs. 6A through 6D are block diagrams of a plurality of repeater modules and selected communications therebetween during various stages of a loop detect process according to an embodiment of the present invention;

Fig. 7 is a flow chart of a process for correcting an isolated node configuration; and

Fig. 8 is a state diagram of a process, according to an embodiment of the present invention, as understood in conjunction with Tables I, II and III.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

Although some or all features of the present invention can be implemented in a wide variety of electrical, optical or electro-optical systems, in one illustrative example, as illustrated in Fig. 2, the nodes of a system can be a plurality of Ethernet repeaters 212a, 212b, 212n coupled to a Ethernet switch 214 in an Ethernet network. Examples of repeaters that can be used in connection with the present invention include those as described in U.S. Patent Application Serial Nos. 09/321,066; 09/330,434; and 09/330,733, *supra*, in which the repeaters are gigabit Ethernet repeaters. In the depicted embodiment, the repeaters 212a,b,n can act as three-port repeaters to provide a shared channel of communication 216 "Port C" between the host 214 and other repeaters coupled to first and second ports 218a,b over cables 222a,b or other external links, or as a short-haul full duplex link between two systems equipped with repeaters. In PCS-bypass mode, a repeater receives data from the host, one of the ports, or one of the external links 222a, 222b and retransmits data to the other link (with timing and signal levels restored). In half-duplex mode, the repeater receives and decodes data from any of its three links. After restoring the signal timing, amplitude and coding, the repeater retransmits the data to the other two links, if no other carrier event is detected. The illustration of Fig. 2 is simplified, at least in that circuitry for providing timing, amplification, and buffering or storage of data is not expressly depicted, although those of skill in the art will understand how to provide and use repeaters in the context of an Ethernet system to implement embodiments of the present invention, at least after understanding the present disclosure.

In the illustration of Fig. 2, connection among the various ports and/or generation of signals to be sent, or handling or storage of signals received, is at least partially controlled by one or more state machines 224a, 224b or other control circuitry 226. A state machine according to one embodiment of the present invention will be described more thoroughly below. In the embodiment of Fig. 2, under control of the state machines 224a,b, the ports 218a,b may be isolated, e.g., by operation of (opening of) switches 228a,b (although shown as mechanical switches, switches 228a,b would typically be provided as transistor or electronic switches).

Preferably, each repeater 212a,b,n has a capability (e.g. implemented in a state machine 224a,b) of detecting whether there is an operative communication link from either or both of the

ports 218a, 218b and another node. In one embodiment, a link is detected if a specified synchronization signal is detected and a full-half duplex auto negotiation process completes. An example of full-half duplex auto negotiation is described, e.g., in Serial No. 09/330,733, *supra*.

According to one embodiment of the invention, a loop detect process is initiated and a repeater detects that a link has been added 312 (Fig. 3). In this embodiment, in response to detection of a link added, the repeater will send a "loop detect" packet out through both of its ports, if both ports have links attached then. In one embodiment, the loop detect packet (which may, in some embodiments, include a plurality of sub-packets) includes a field for at least partially specifying the address of the node 314. In one embodiment, each repeater uses the media access controller (MAC) address of the gigabit port on the Ethernet switch it is connected to, as its own address.

Although any number of protocols could be devised and used for sending "loop detect" messages, in one embodiment, the "loop detect" messages (and, preferably, other messages used in the system as described herein) is similar in form and protocol to messages already used in a full/half duplex auto-negotiation system. Preferably, the auto-negotiation system is point-to-point, such that a "loop detect" message, sent to a neighbor, will not be received by other nodes in the system unless the neighbor re-generates it. An advantage of using a form or protocol for messages similar to that already used (e.g., in an installed base of apparatus) is to facilitate implementing invention in a fashion which can be backwards compatible (so that implementation of the present invention does not disrupt or otherwise substantially affect operation of current apparatus) and/or such that little or no modification or replacement of existing apparatuses is required in order to implement the present invention. In one embodiment, the present invention can be implemented by using existing Ethernet switches (or other components) in connection with repeaters which have been configured with state machines 224ab which implement embodiments of the present invention, e.g., as described more thoroughly below. In one embodiment, a "loop detect" message is substantially similar in length to an auto negotiate message; but, in a field which, in the auto negotiate message, always has a first content (e.g. all zeros), the "loop detect" message will have a different content (at least one non-zero bit) to identify it as other than an auto negotiate message and, in particular, as a "loop detect" message.

As will be understood by those of skill in the art, the repeaters can include circuitry which parses this field of messages and handles the message as an auto negotiate message if this field contains all zeros and, otherwise, handles this message (e.g. as described below) as a "loop detect" message in response to other contents in this field. Typically, the "loop detect" message will contain additional information such as a "least MAC address." In general, in a point-to-point system a source address field and/or a destination address field is not needed or provided. If desired, data integrity can be provided by an acknowledgment field which echos the data field(s) previously received.

As depicted in Fig. 4, when any repeater receives a loop detect packet, which will contain some address (designated "K") as the least MAC address, 412, the repeater will compare the value of K to its own node address 414. If the repeater determines that K is greater than its own address, the packet will generate a "loop detect" message, which will contain its own address as the "least MAC address", thus creating a new value for "K" and will then send the packet out through its opposite port, i.e. the port other than that where the packet was received (at least, if the opposite port is coupled to another node). It is noted that, at least in this embodiment, although two "least MAC addresses" can be sent out (from the two ports) substantially simultaneously, each is sent out from the port which is opposite the port where it was received. That is, a received message is not sent directly back to the repeater where it was just received from. If the repeater determines that K is less than the repeater's own address, the repeater will send the "loop detect" packet out through its opposite port unaltered, i.e., containing K, as received, as the "least MAC address" 424. If the repeater determines that the value of the received "least MAC address" (i.e. K) is equal to its own address 426 that repeater will set a flag indicating "master loop breaker" (MLB) status 428. If two (neighboring) repeaters both determine, at substantially the same time, that there is a loop, preferably, a tie-breaker protocol is provided, such as allowing the repeater which has the larger address to declare itself the maser loop breaker. Preferably, each repeater has information indicating the MAC address of its neighbor(s).

At this point, the presence of a loop has been detected and it would be possible to implement embodiments of the invention in which the process stops at this point and/or outputs

an indication of the existence of a loop, e.g. for appropriate manual handling by an operator and the like. Preferably, however, in response to the detection of a loop, the MLB will disable one of the ports, in this example, it's own second or B port 432. Preferably the port is disabled in a fashion such that it can later be reenabled. In some embodiments, disabling the port involves suspending receipt or transmission of normal communication packets through the port (although, in some embodiments, the port may still be used for receiving or transmitting negotiation or similar system or control packets). In this way a communication link 432 can be effectively suspended or eliminated, even without physically removing or switching off the link medium.

Figs. 5A through 5G depict one example of a series of communications among repeaters resulting in detection of a loop, according to an embodiment of the present invention. In the example of Figs. 5A through 5G, four repeaters 512a,b,c,d (coupled to a Ethernet switch, not shown) each contain first or "A" ports 516a,b,c,d and second or "B" ports 518a,b,c,d. In each instance, a B port 518 of a repeater is coupled to an A port of a neighbor repeater by a communication link 514a,b,c,d. Each repeater has a node address and includes a memory or other device 522a,b,c,d for storing (or receiving) its own address. In the example of Fig. 5A, the first repeater 512a has an address of two 522a, the second repeater 512b has an address of one 522b, the third repeater 512c has an address of three 522c and the fourth repeater 512a has an address of zero 522d. As an illustrative example, if the configuration of the system is changed, e.g., to add the last communication link 514d, achieving the configuration depicted in Fig. 5A, when a repeater, e.g., 512a, detects the occurrence of a new link 514d, it outputs from it's A and B ports 516a, 518a, "loop detect" packets 524a, 525a each of which includes a field storing a value K 526a, 527a equal to the repeater's 512a own address 522a, in this case K=2. When the second repeater 512b receives this packet 524, using a procedure similar to that depicted in Fig. 4, it will compare the value of K 526a to its own address 522b. In this case, K is greater than its own address 416 and accordingly, the second repeater 512b, as depicted in Fig. 5B, will output a loop detect packet 524b which will be similar to the received packet except that the "least MAC address" field will have a value equal to the address 522b of the second repeater 512b. Similarly, repeater 512d will output from it's a port a packet 525b having a least MAC address with 527b with a value of K = 0 (since repeater 512d's own address 522d is zero and is thus less

than the received least MAC address 527a). As depicted in Fig. 5C, when the third repeater 512c compares its own address 522c to the "least MAC address" of the received packet 524b, since its own address is greater than the value of K in the received packet 422, the third repeater will output a packet 524c which will be substantially identical to the received packet, i.e., which will have the same value in the "least MAC address" field 526c as was contained in the corresponding field of the received packet 526b. At substantially the same time, the third repeater 512c compares its own address 522c to the "least MAC address" of the received packet 525b. Since its own address is greater than the value of K in the received packet, the third repeater will also output, from it's A port, a packet 525c which will be substantially identical to the received packet, i.e., which will have the same value in the "least MAC address" field 527c as was contained in the corresponding field of the received packet 527b.

The fourth repeater 512d, having an address 522d less than the "least MAC address" of the received packet 526c, will output a packet 524d (Fig 5D) which has its own address 522d (in this case a value of zero) in the "least MAC address" field 526c (i.e., K=0). At generally the same time, the second repeater 512b, having an address 522b greater than the "least MAC address" of the received packet 527c, will output a packet 525d which has the same value in the "least MAC address" field 527d as was contained in the corresponding field of the received packet 527c.

As depicted in Fig. 5E the first repeaters 512a, having an address greater than the received "least MAC address" 527d will output a loop detect packet 525e from it's A port, to the fourth detector 512d which has the same value in the "least MAC address" field 527e as was contained in the corresponding field of the received packet 527d.. In the configuration of Fig. 5E, the last or fourth repeater 512d will thus receive a packet 525d containing a "least MAC address" 527e which is equal to its own address 522d. Accordingly, using the procedure of Fig. 4, the fourth repeater 512d will set a flag declaring itself to be the master loop breaker, and will set a switch to isolate its B port 518d, thus effectively breaking the loop.

Figs. 6A through 6C depict another procedure for detecting a loop. The procedure of Fig. 6A through 6C is similar to that of Figs. 3 through 5 except that a repeater will declare the existence of a loop if it receives two packets (at it's A and B ports) which have the same "least

MAC address". Thus, in the embodiment of Fig. 6A, when the first repeater 512a detects a link added 412, it outputs a first loop detect packet 524a (as described above) through the first link 514a, (which will be received by the second repeater 512b) and outputs an identical loop detect message 624a through it's A port 516a, (which will be received by the fourth repeater 512d). As depicted in Fig. 6B, in response, the second repeater 512b outputs a loop detect packet 524b (having a "least MAC address equal to 1) and, substantially simultaneously the fourth repeater 512d outputs, through it's a port 516d a loop detect packet 624b having its least MAC address 626b equal to zero (since the address 522d of the fourth repeater 512d is less than the least MAC address of the received packet 624a). Thus, as seen in Fig. 6B, the third repeater 512c will receive two loop detect packets 524b, 624b. As seen in Fig. 6C, the fourth repeater 512d will receive a packet 624d which has, as its least MAC address 626d a value equal to one, and the second repeater 512b will receive a packet 624c which has, as its least MAC address 626c, a value equal to zero. As seen in Fig. 6D, the first repeater 512A will receive two packets, respectively, through it's a and B ports. The two received packets 624e, 624f have identical values for the "least MAC address," 626e,f, namely a value of zero. In response, the first repeater 512a will set a flag indicating its master loop breaker status 428 and will set a switch to isolate its B port 518a.

As can be seen from comparing Figs. 5A through 5E with Figs. 6A through 6D, a potential advantage of the procedure of Figs. 6A through 6D is that it may complete a loop detect process in fewer steps or cycles, compared to the process of Figs. 5A-E.

As depicted in Fig. 1C, once a communication link 114a of a loop configuration has been effectively disabled, the remaining configuration provides an open or non-loop system in which all nodes can at least indirectly communicate with each other. However, if, subsequently, another communication link becomes disabled, as depicted in Fig. 1D, 114b one or more nodes 112b may become isolated, i.e., without any communication link which can be used for communicating with at least some of the remaining or other nodes. In the embodiment of Fig. 7, in response to the detection of a lost link 712 (such as the effective disabling of link 144b) a lost link message 712 will be sent. In the embodiment of Fig. 7, each repeater which receives a "lost link" message, if it is not the MLB 714, will pass the "lost link" message to the next node 716.

When the "lost link" message is received by the MLB node, it will reactivate its B port 718 (i.e. port B_1) in the embodiment of Fig. 6E), causing the effective reactivation of the associated communication link 114a, thus restoring the system to a configuration in which all nodes can communicate with one another, i.e., in which there are no isolated nodes.

The MLB will then unset its MLB flag 722. The reactivation of a previously (intentionally) deactivated communication link can also be achieved using procedures other than that depicted in Fig. 7. For example, in one procedure, each node which receives a "lost link" message will assure that both its ports are active and will make sure that its MLB flag is unset and will then pass on the "lost link" message to neighboring nodes. Thus, in this procedure there is no need for each node to specifically determine whether it is the MLB and to take different actions depending on the result.

In general, it can be advantageous to provide a system having a configuration similar to that depicted in Fig. 1C (in which there is a "redundant" deactivated link 114a which, if activated, would create a loop configuration), since this affords the opportunity to reconfigure a system which has lost a link (as depicted in Fig. 1D) so as to restore desired communication abilities (as depicted in Fig. 1E). Preferably, and using procedures similar to those described above, any or all of the detecting of a loop, the breaking of a loop or the reactivation of a link in response to a lost link can be achieved substantially automatically, i.e. without the need for a user to note or respond to light or other signals or outputs, and without the need to manipulate or reconfigure cabling or otherwise perform some manipulation. However, even though the present invention can be implemented fully or partially automatically, i.e. without the need for human intervention or manipulation, nevertheless, if desired, certain human involvement may be provided for, such as by providing for LED or other light output, computer console output and similar output indicating the presence of a loop detection, a loop breaking and/or a communication link reactivation. Such output may be useful in system maintenance, troubleshooting and the like.

Fig. 8 depicts a state diagram that can be used, e.g., in implementing a state machine according to an embodiment of the present invention. Those of skill in the art will understand how to provide a state machine to implement a state diagram, e.g., as depicted in Fig. 8, at least

after understanding the present disclosure. As depicted in Fig. 8, a state machine can be in any of a plurality of different states 812a-h. Table I provides a brief description of states 812a through 812h. Transitions from one state to another are triggered by various events indicated in Fig. 8 by numerals 1 through 22. Table II describes events 1 through 22. Table III is a state transition action table for the state diagram of Fig. 8.

TABLE I

The states shown in Fig. 8.

A. Sync Lost State: The sync on the link is lost. It will transition into state B when the sync comes back.

B. Exchange Self Mac Address State: Send the node's own MAC address to the link's neighbor, and wait for acknowledgment. It will transition to state C or D after the acknowledgment for the last byte of the MAC address is received.

C. MLB Selection: Send the local least MAC to its link neighbor, and compare the received MAC with the local least MAC to see if it is equal to its own MAC to decide if it should appoint itself as the Master Loop Breaker.

D. No Loop State: One other link in the node stack has lost sync. It could be the other link on the same node, or the other link on the same node received Sync_Lost_Detected code word from its neighbor. It will move to state F state after Sync_Lost_Detected is sent to the neighbor (an acknowledgment is received) or to state C when the syncOK is detected from the other link

TABLE II

Events that trigger the state transitions for Fig. 8.

1. SyncOK detected.

2. SyncOK Lost Detected.

3. Lpbr_Start code word received.

4. Sync_OK_Detected code word received from the link neighbor.

5. Sync_Lost_Detected code word received from the link neighbor.

6. Link_Enable code word received form the link neighbor.

7. MAC ID byte received from the link neighbor and both external links are sync OK.

8. MAC ID byte received from the link neighbor and only one link has sync OK.

9. Sync_OK_Detected code word received on the other link (or Sync OK detected on the other link.

10. Sync_Lost_Detected code word received on the other link (or Sync lost on the other link).

11. Linc_Enable code word received on the other link.

12. MAC ID received on the other link and it is equal to the self MAC sent form the link (MLB chosen to be the node itself).

13. Ack of Sync_OK_Detected code word received.

14. Ack of Sync_Lost_Detected cord word received.

15. Ack of Link_Enable code word received.

16. Ack of MAC ID byte received.

17. Link neighbor becomes normal (code word = 0).

18. Node is in the process of changing duplex mode.

19. Timeout: No acknowledgment received after 15 seconds.

20. Autonegotiation results in full-duplex mode.

21. Duplex change in progress.

22. Duplex change in progress.

The following is a table showing actions taken upon each state transition: (X means any state from A to G except the end state)

TABLE III

State Transition	Event	Actions
A -> B	1	Send Lpbr_Start code word to the link neighbor, activate the timer event that will wake up high priority process periodically.
B -> D	8	Send the Sync_Lost_Detected to the link neighbor.
B -> B	16	Send the next byte of the port's own MAC Address ID.
B -> C	7	Send the first byte of the local least MAC to the link neighbor.
B -> F	19	Turn on Tx/Rx on the link, deactivate the timer event.
C -> D	10	Send the Sync_Lost_Detected to the link neighbor.
C -> C	13, 16	Send the next byte of the least MAC address ID to the link neighbor. Compare the received least MAC with the port's own MAC.
C -> E	12	Send Link. Enable code word to the neighbor. Turn off Tx/Rx on the redundant link.
C -> E	11	Send Link_Enable code word to the link neighbor.
D -> C	9	Send Sync_OK_Detected to the neighbor link.
B, D -> F	14	Enable Tx/Rx on the link, deactivate the timer event.
E -> D	4	Send the Sync_Lost_Detected to the link neighbor.
F -> G	5, 6	Disable C link, activate timer.
F -> C	4	Send Sync_OK to the link neighbor.
G -> A	14	Toggle the link to reset state on the link neighbor and itself.
G -> F	14, 15, 17	Turn on Tx/Rx on the link, deactivate the timer event.
H -> H	22	Take necessary steps to change duplex mode.

State Transition	Event	Actions
B, C, E, F, G -> D	10	Send the Sync_Lost_Detected to the link neighbor. Turn Tx/Rx of the redundant link on if the node is the MLB and Tx/Rx on the redundant link is off.
C, D, E -> B	19	Restart the loop breaking process, activate the timer if not already active.
F, H -> B	3	Restart the loop breaking process, activate the timer if not already active.
B, C, D, E, F, G -> A	2	Reset state machine, deactivate the timer event.

Although Fig. 8 and Tables I through III provide an example of a system which can operate according to an embodiment of the present invention, the present invention can also be implemented in other fashions such as using other configurations of state machines, or configurations which do not use traditional state machines, such as implementing the invention by control provided by a programmed microprocessor or computer.

In light of the above disclosure, a number of advantages of the present invention can be seen. The present invention can provide a system for detecting the presence of a loop in an electrical, optical or electro-optical system having multiple nodes and communication links, preferably in a substantially automatic fashion. The present invention can provide for breaking a loop, preferably substantially automatically, so as to provide a open-type topology which still permits all nodes to at least indirectly communicate with all other nodes. The present invention can provide a system for reconfiguring the system to convert it from a system in which some nodes are isolated to a system in which all nodes can communicate with one another, at least indirectly. In at least some embodiments, the present invention can be implemented in a fashion which is substantially backwards compatible with existing protocols of the installed base of apparatus and/or which requires little or no modification or replacement of existing apparatus.

A number of variations and modifications of the invention can be used. It is possible to use some features of the invention without using others. For example, it is possible to use loop detecting without loop breaking. It is possible to use reconfiguration to correct isolated nodes

without using loop breaking or detecting. Although the invention has been described in connection with a gigabit Ethernet network, some or all features of the present invention can be used in connection with other systems including other types of computer-based networks, local area networks, wide area networks, Internet installations or components, voice communication systems such as land line, microwave, cellular or satellite-based voice communications and the like. Although the present invention can be implemented in a substantially modular-repeater environment including as described in Serial No. 09/330,733, *supra*, some or all features of the present invention can be used in connection with substantially integrated or otherwise non-modular repeaters or other components. Although the present invention was described including by examples in which nodes were repeaters of switch components, some or all features of the present invention can be implemented and systems where nodes are other types of devices including hubs, routers, switches, bridges, gateways, personal computers, printers or other peripheral devices, telephones or other communication devices and the like. Although embodiments were described in which packets are sent from both ports of a repeater, it is also possible (although not necessarily desirable) to implement operable embodiments in which packets are output only through one port, as part of a loop detect procedure.

The present invention, in various embodiments, includes components, methods, processes, systems and/or apparatus substantially as depicted and described herein, including various embodiments, subcombinations, and subsets thereof. Those of skill in the art will understand how to make and use the present invention after understanding the present disclosure. The present invention, in various embodiments, includes providing devices and processes in the absence of items not depicted and/or described herein or in various embodiments hereof, including in the absence of such items as may have been used in previous devices or processes, e.g. for improving performance, achieving ease and\or reducing cost of implementation. The present invention includes items which are novel, and terminology adapted from previous and/or analogous technologies, for convenience in describing novel items or processes, do not necessarily retain all aspects of conventional usage of such terminology.

The foregoing discussion of the invention has been presented for purposes of illustration and description. The foregoing is not intended to limit the invention to the form or forms

disclosed herein. Although the description of the invention has included description of one or more embodiments and certain variations and modifications, other variations and modifications are within the scope of the invention, e.g. as may be within the skill and knowledge of those in the art, after understanding the present disclosure. It is intended to obtain rights which include alternative embodiments to the extent permitted, including alternate, interchangeable and/or equivalent structures, functions, ranges or steps to those claimed, whether or not such alternate, interchangeable and/or equivalent structures, functions, ranges or steps are disclosed herein, and without intending to publicly dedicate any patentable subject matter.

What is claimed is:

1. A method for reconfiguring a communication system comprising a plurality of nodes coupled by a plurality of communication links comprising:

determining that said communication system includes a closed loop topology in response to receipt of a communication at at least a first of said plurality of nodes;

at least temporarily preventing effective communication across a selected one of said plurality of communication links to change said closed loop topology to an open topology.

2. A method as claimed in claim 1 further comprising designating one of said plurality of nodes as a loop-breaking master in response to said step of determining, wherein said step of preventing is performed in response to a control signal or communication output by said loop-breaking master.

3. A method as claimed in claim 2 wherein said first node is said loop-breaking master.

4. A method as claimed in claim 1 further comprising reestablishing effective communication over said selected one of said plurality of communication links in response to detection of a link loss.

5. A method for detecting loop topology in a communication network having a plurality of nodes coupled by a plurality of links wherein each of said plurality of nodes is associated with a determinable node value comprising:

sending at least a first communication from a first node to at least a second node, said first communication including an indication of said node value of said first node;

said second node receiving said communication from said first node which includes a received node value and comparing said received node value to a first node value which is the

node value of said second node, said second node outputting a signal indicative of a closed loop topology when said received node value equals said first node value.

6. A method as claimed in claim 5 further comprising said second node outputting a communication which includes said received node value when said received node value is closer to a predetermined node value than said first node value.

7. A method as claimed in claim 5 further comprising said second node outputting a communication which includes said first node value when said first node value is closer to a predetermined node value than said received node value.

8. A method as claimed in claim 5 further comprising said second node outputting a communication which includes said received node value when said first node value is greater than said received node value.

9. A method as claimed in claim 5 further comprising said second node outputting a communication which includes said first node value when said first node value is less than said received node value.

10. A method as claimed in claim 6 wherein said node values are node addresses.

11. A method as claimed in claim 1 wherein said nodes include ethernet repeaters coupled to an ethernet switch in an ethernet network.

12. A method as claimed in claim 2 wherein said selected one of said plurality of communications links communication link is a communication link coupled to said loop-breaking master.

13. A method as claimed in claim 5 wherein said step of sending said first communication from said first node is performed in response to a change in the number of nodes or links in the system.

14. A method for avoiding node isolation in a network communication system having a plurality of nodes coupled by a plurality of communication links, the method comprising:

deactivating at least a first communication link to provide a system having an open topology with no isolated nodes;

detecting effective loss of a communication link; and

reactivating said first communication link.

15. A method as claimed in claim 14 wherein said step of deactivating is performed in response to detection of a closed loop topology.

16. Apparatus for reconfiguring a communication system comprising a plurality of nodes coupled by a plurality of communication links comprising:

a state machine in at least one of said plurality of nodes configured to determine that said communication system includes a closed loop topology in response to receipt of a communication at said one of said plurality of nodes;

said state machine also configured to provide a control signal to at least temporarily prevent effective communication across a selected one of said plurality of communication links to change said closed loop topology to an open topology.

17. Apparatus as claimed in claim 16 wherein said state machine is further configured to designate one of said plurality of nodes as a loop-breaking master in response to said step of determining, wherein said control signal or communication output by said loop-breaking master.

18. Apparatus as claimed in claim 16 wherein said state machine is further configured to reestablish effective communication over said selected one of said plurality of communication links in response to detection of a link loss.

19. Apparatus for avoiding node isolation in a network communication system having a plurality of nodes coupled by a plurality of communication links, the apparatus comprising:

means for deactivating at least a first communication link to provide a system having an open topology with no isolated nodes;

means for detecting effective loss of a communication link; and

means for reactivating said first communication link.

20. Apparatus as claimed in claim 19 wherein said means for deactivating, means for detecting and means for reactivating comprises at least a first state machine coupled to at least a first of said plurality of nodes.

21. A method as claimed in claim 19 wherein said means for deactivating includes means for deactivating in response to detection of a closed loop topology.

ABSTRACT

In a system having a plurality of nodes such as Ethernet repeaters, coupled by communication links such as cables, systems and protocols are provided for detecting and /or breaking loops. In one aspect, in response to an added link, a repeater sends a "detect loop" message containing its address to at least one neighbor. Each repeater which receives the "detect loop" message, in turn, sends it to its own neighbor, with the lesser of the received address and its own address. A repeater which receives a "detect loop" message containing its own address declares itself a master loop breaker and can isolate one of its ports to break the loop. In one aspect, a previously intentionally-isolated port can be re-activated, e.g., in response to the loss of a communication link which could potentially isolate one or more nodes.

M \4103\-61641\APP-patent app wpd

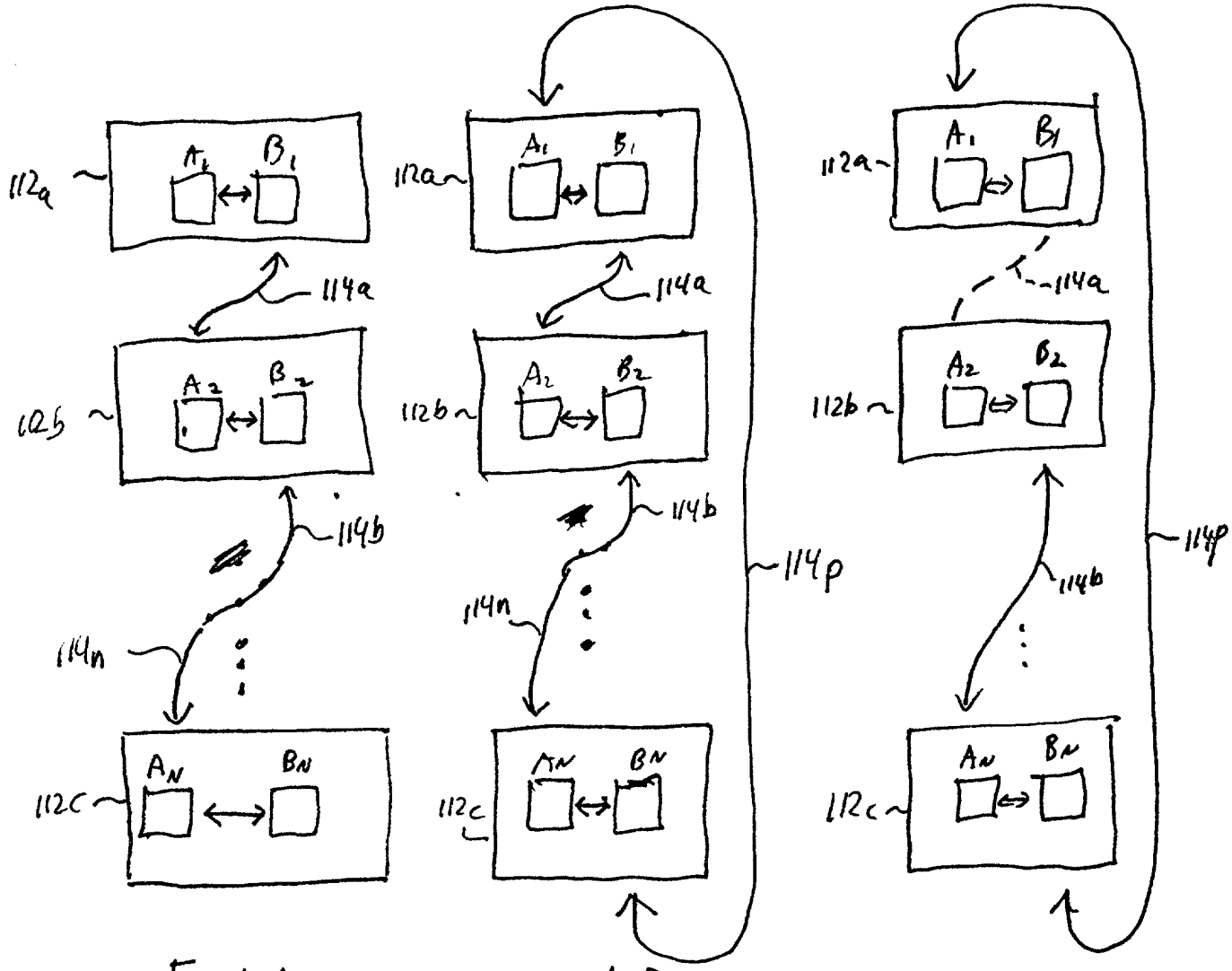


Fig 1A

Fig 1B

Fig 1C

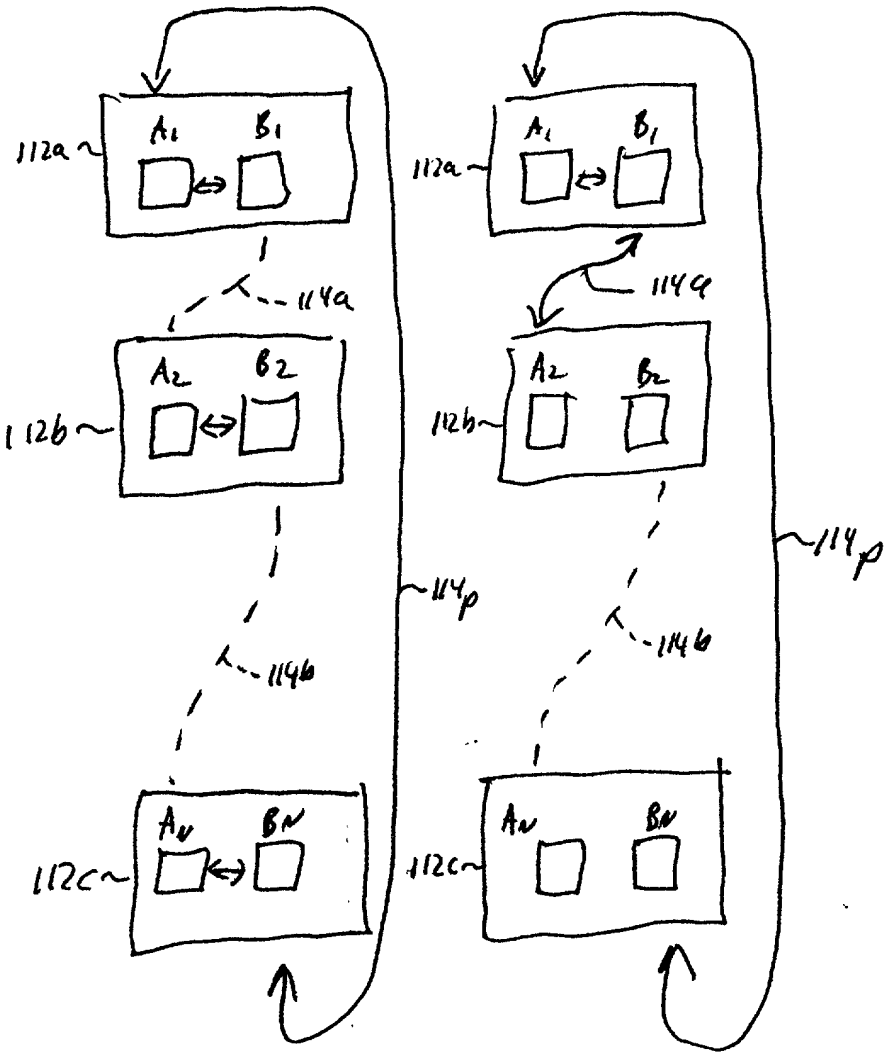
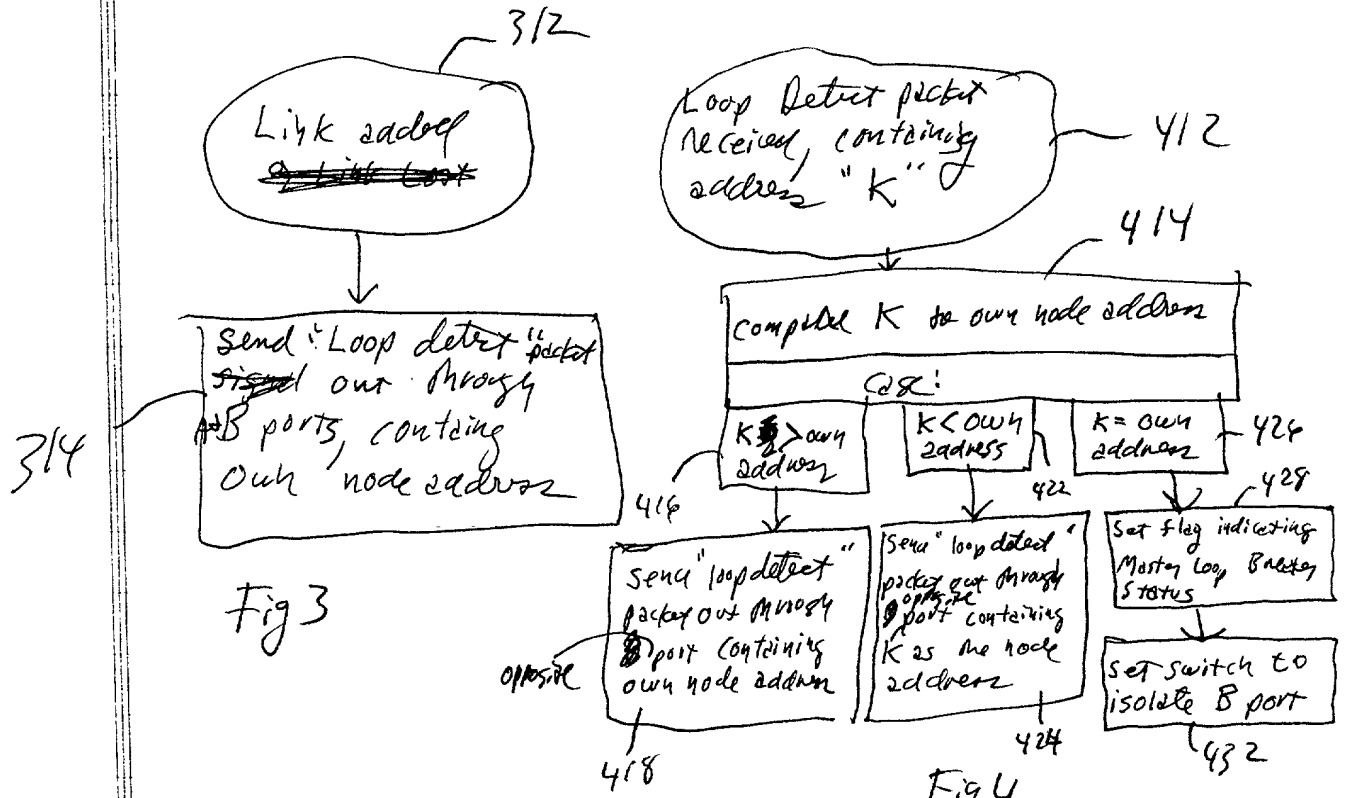
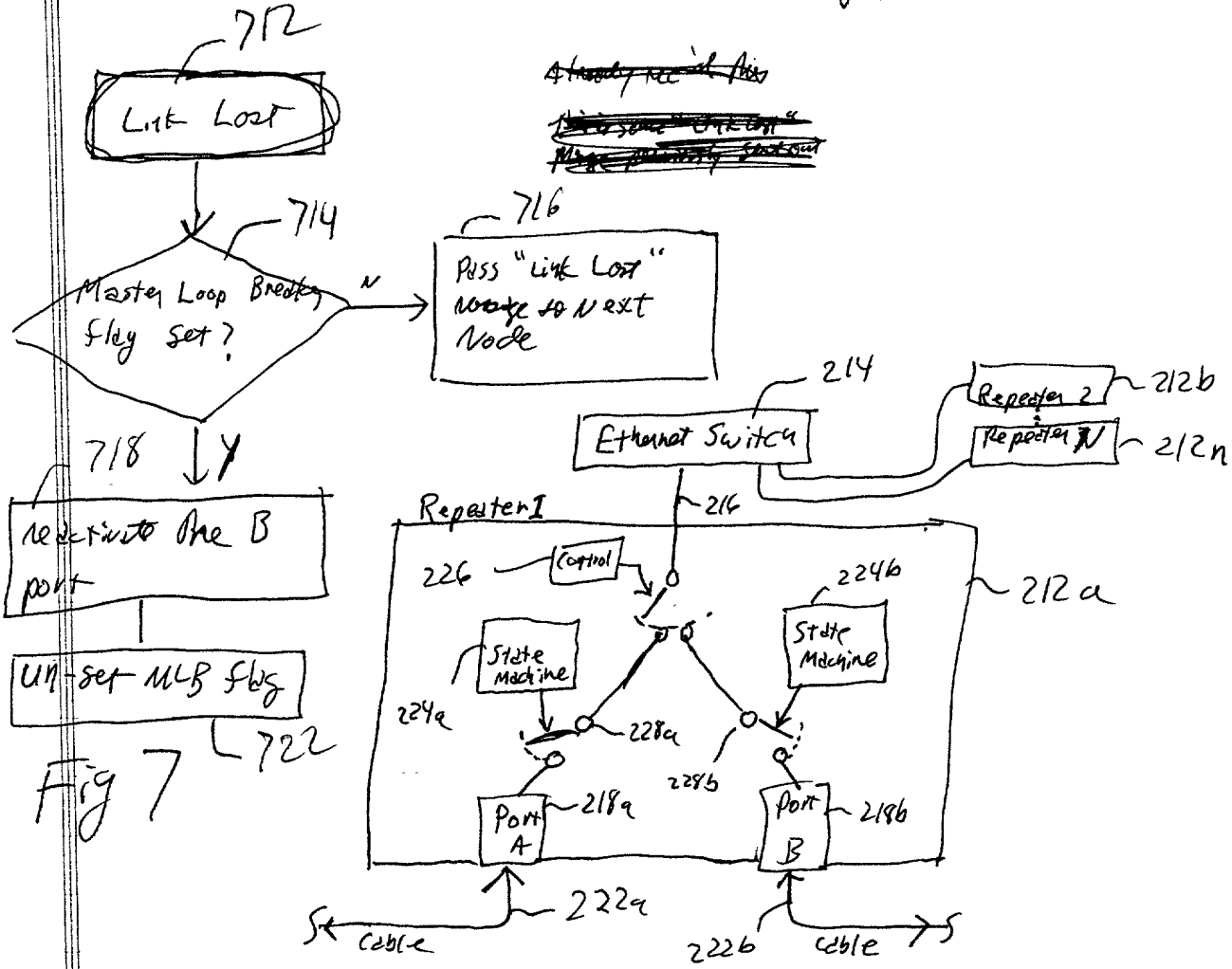


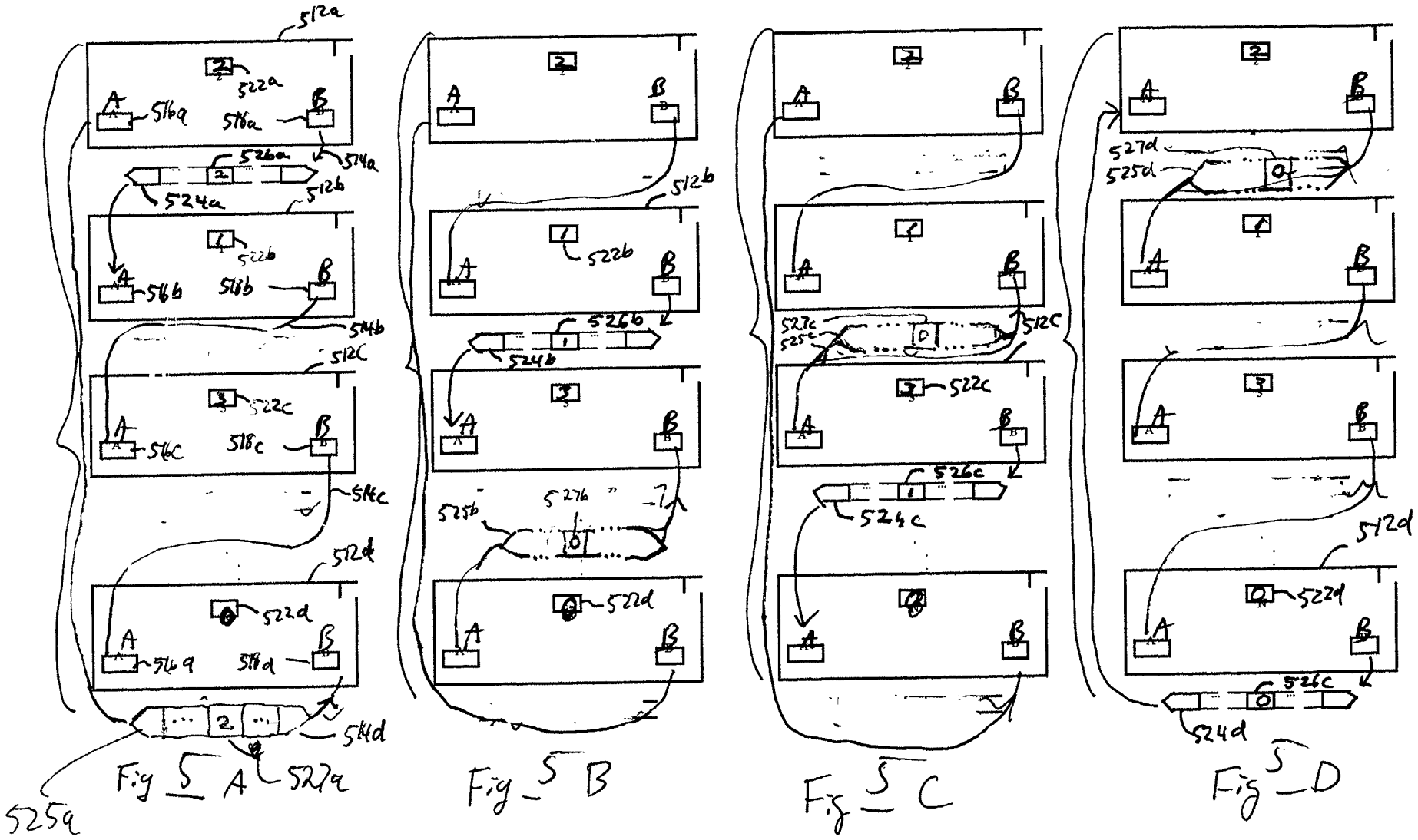
Fig 1D

Fig 1E


312
Link added
Send "Loop detect" packet out through A+B ports, containing own node address
314
Fig 3
Loop Detect packet received, containing address "K"
412
414
compare K to own node address
case:
K > own address
K < own address
K = own address
416
422
424
428
Send "loop detect" packet out through opposite port containing own node address
Send "loop detect" packet out through opposite port containing K as the node address
Set flag indicating Master Loop Breaker Status
Set switch to isolate B port
418
424
432
Fig 4


712
Link Lost
714
Master Loop Breaker flag set?
N
716
Pass "Link Lost" message to Next Node
Y
718
redirect to the B port
un-set MLB flag
722
Fig 7
214
Ethernet Switch
Repeater 2
212b
Repeater N
212n
Repeater 1
216
226
Control
224a
224b
State Machine
State Machine
228a
228b
212a
Port A
218a
Port B
218b
222a
222b
cable
cable
Fig 2

DOCKET GATE960



C:\Temp1\cicco Fig. 2.wpd

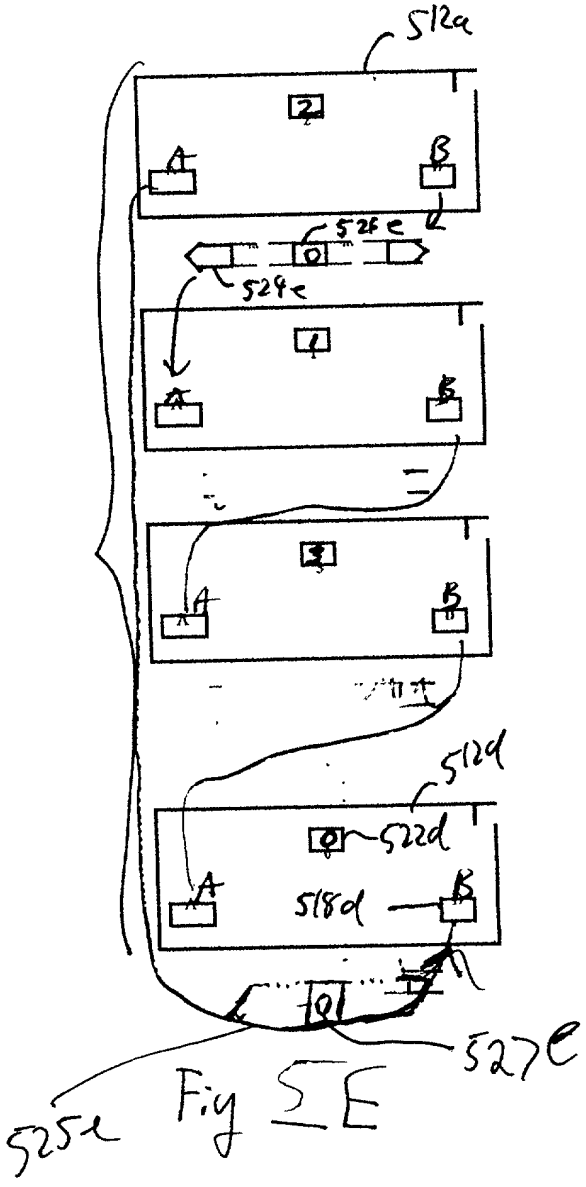


Fig 5E

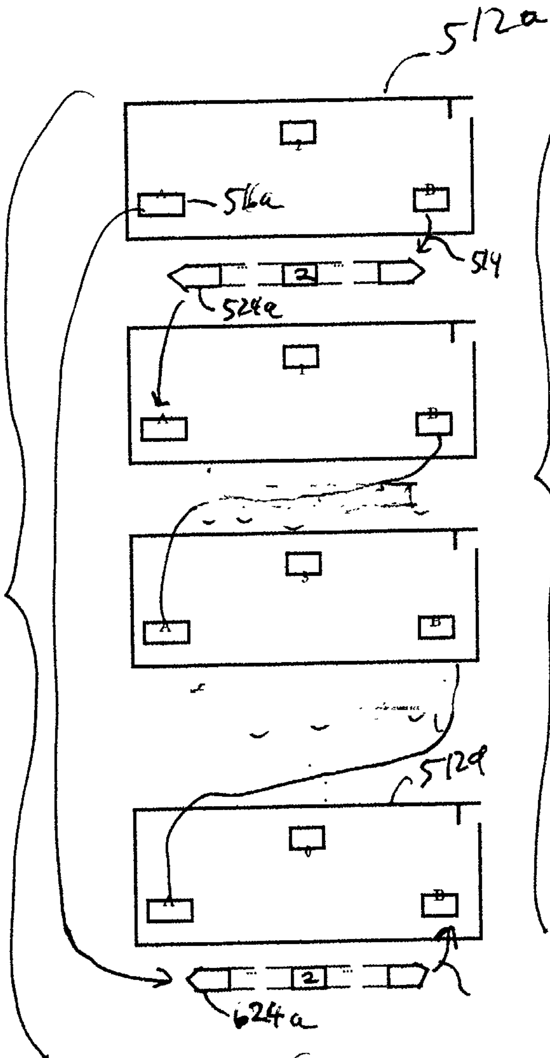


Fig 6 A

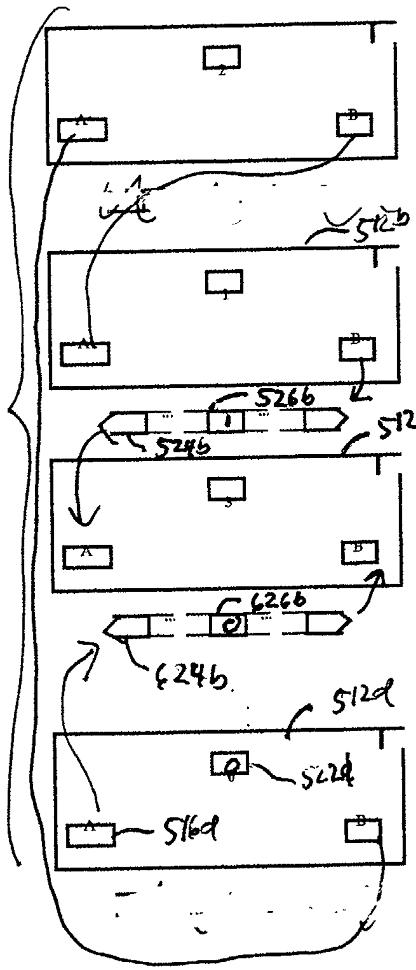


Fig 6 B

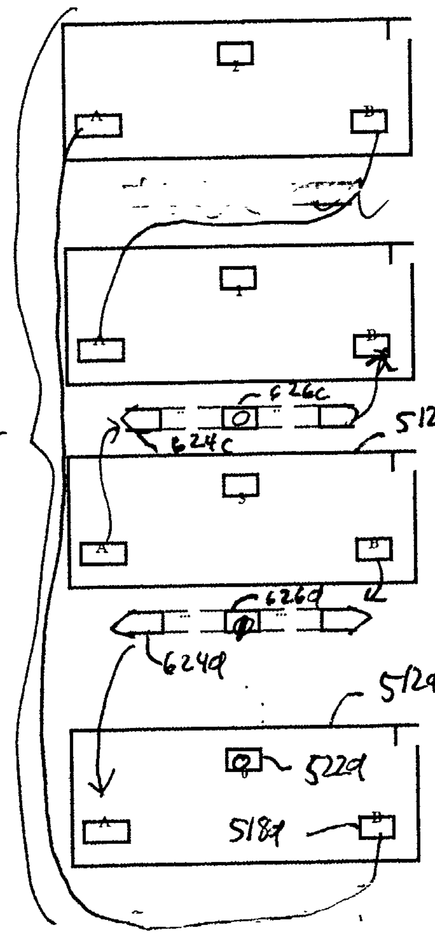


Fig 6 C

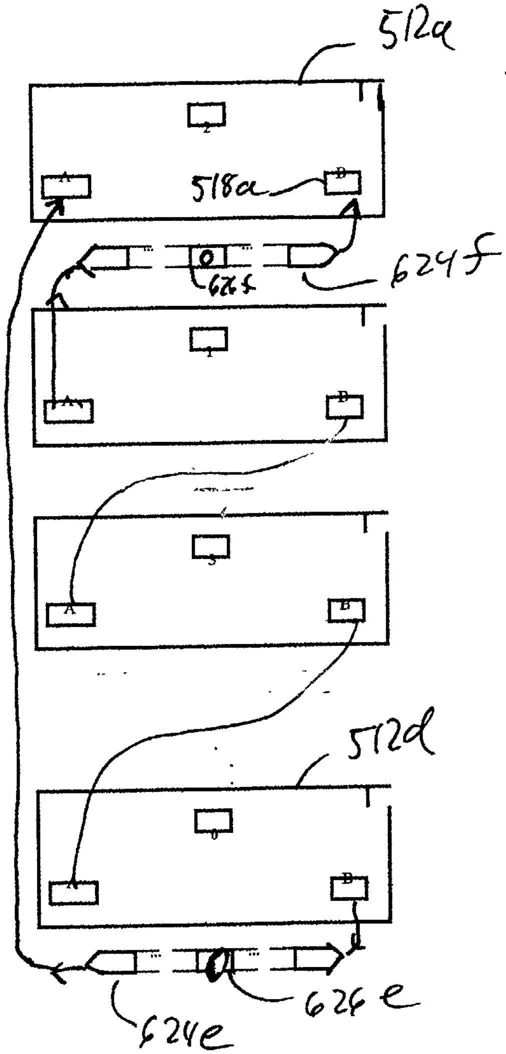


Fig 6 D

C:\Temp1\cicco Fig. 2.wpd

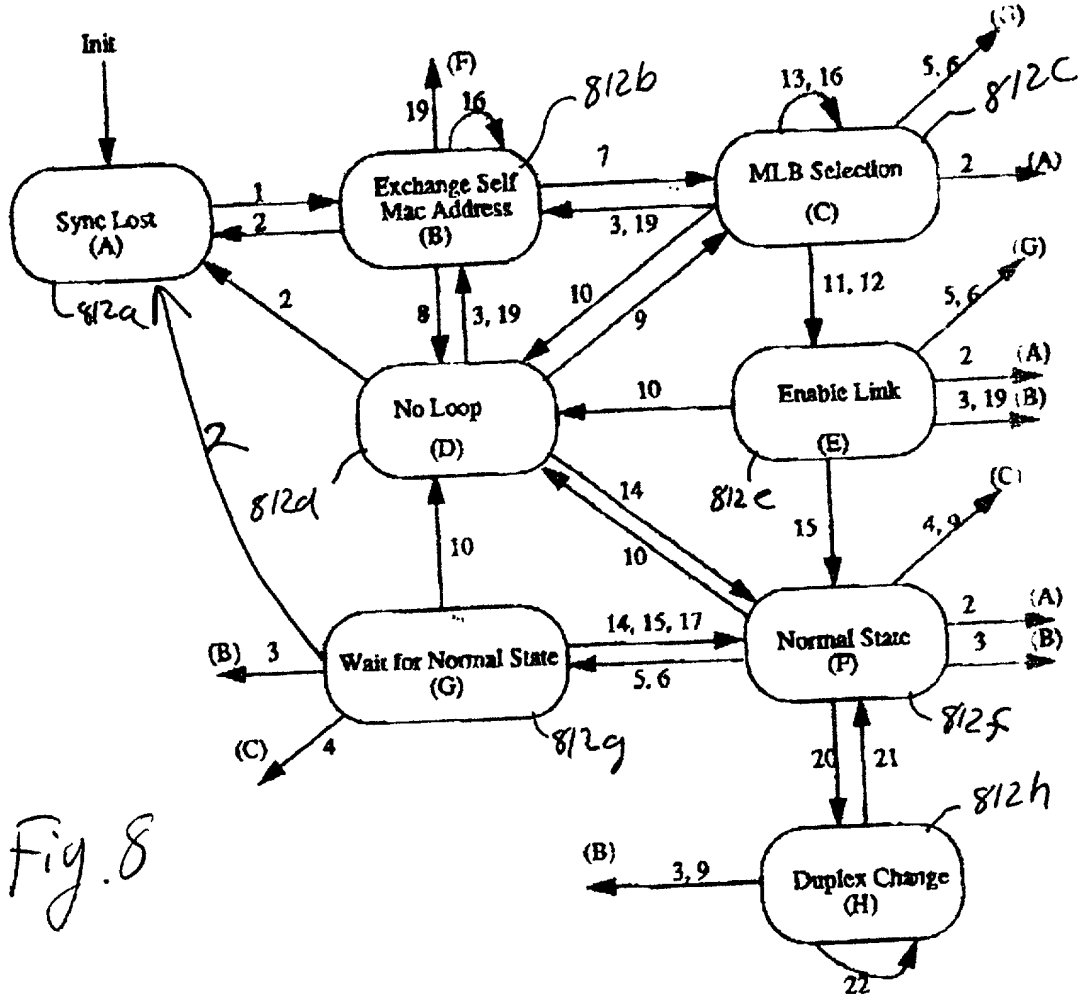

Init
Sync Lost
(A)
Exchange Self
Mac Address
(B)
MLB Selection
(C)
No Loop
(D)
Enable Link
(E)
Normal State
(F)
Wait for Normal State
(G)
Duplex Change
(H)
812a
812b
812c
812d
812e
812f
812g
812h
1
2
3, 19
7
8
9
10
11, 12
13, 16
14
15
14, 15, 17
5, 6
4, 9
20
21
22
3, 9

Fig. 8